ATGAAGATTACAAAACCACATGTGGCCATGTTCGCTAGCCCCGGAATGGGCCACA TCATCCCGGTGATCGAGCTCGGAAAACGCTTAGCTGGTTCCCACGGCTTCGATGT CACCATTITCGTCCTTGAAACCGACGCAGCCTCAGCTCAATCTCAATTCCTTAACT CACCAGGCTGCGACGCCCTTGTTGATATCGTTGGCCTCCCAACGCCCGATAT CTCCGGTTTAGTCGACCCATCAGCCTTTTTTGGGATCAAGCTCTTGGTCATGATGC GTGAGACCATTCCTACCATCCGGTCAAAGATAGAGGAGATGCAACACAAACCAA CGGCTCTGATCGTAGACTTGTTTGGTTTGGACGCGATACCGCTCGGTGGTGAGTTC AACATGTTGACTTATATCTTCATCGCTTCAAACGCACGTTTTCTCGCGGTGGCTTT GTTTTTCCCAACGTTGGACAAAGACATGGAAGAAGAGCACATAATCAAGAAGCA ACCTATGGTTATGCCTGGATGTGAACCGGTTCGGTTTGAAGATACACTTGAAACA TTCCTTGACCCAAACAGCCAACTCTACCGGGAATTTGTTCCTTTCGGTTCGGTTTT CCCAACGTGTGATGGTATTATTGTGAATACATGGGATGATATGGAGCCCAAAACT TTGAAATCTCTTCAAGACCCAAAGCTCTTGGGTCGAATTGCTGGTGTACCGGTTTA TCCAATTGGTCCTTTGTCTAGACCGGTTGATCCATCTAAAACTAATCATCCGGTTT TGGATTGGTTAAACAAACAGCCGGACGAGTCGGTACTTTACATTTCATTTGGAAG CGGTGGCTCTCTCGGCTAAACAACTAACCGAATTGGCTTGGGGACTTGAGATG AGTCAGCAACGGTTCGTTTGGGTGGTTCGACCCCCGGTGGACGGTTCAGCTTGCA GTGCATATTIATCCGCTAACAGTGGTAAAATACGAGACGGTACACCTGATTATCT CCCGGAAGGTTTTGTTAGCCGGACTCATGAGAGAGGCTTTATGGTCTCTTCTTGG GCTCCCCAAGCGGAGATCTTGGCCCACCAAGCCGTAGGTGGGTTTCTAACTCACT GCGGTTGGAATTCGAGTCGTCGTTGGTGGCGTTCCGATGATCGCGTG GCCACTTTTTGCGGAGCAGATGATGAACGCGACACTCCTCAACGAAGAGCTTGGC GTTGCCGTCCGCTCTAAGAAACTACCGTCGGAGGGAGTGATTACGAGGGCGGAG ATCGAGGCGTTGGTGAGAAAGATCATGGTGGAGGAGGAAGGTGCTGAGATGAGA AAGAAGATAAAGAAGCTGAAAGAGACCGCTGCCGAATCGCTGAGTTGCGACGGT GGAGTGGCGCATGAATCGTTGTCAAGAATCGCCGACGAGCAGCGAGCATCTTTTGG AGCGTGTCAGGTGCATGGCACGTGGTGCCTAG

MKITKPHVAMFASPGMGHIIPVIELGKRLAGSHGFDVTIFVLETDAASAQSQF LNSPGCDAALVDIVGLPTPDISGLVDPSAFFGIKLLVMMRETIPTIRSKIEEMQH KPTALIVDLFGLDAIPLGGEFNMLTYIFIASNARFLAVALFFPTLDKDMEEEHII KKQPMVMPGCEPVRFEDTLETFLDPNSQLYREFVPFGSVFPTCDGIIVNTWDD MEPKTLKSLQDPKLLGRIAGVPVYPIGPLSRPVDPSKTNHPVLDWLNKQPDES VLYISFGSGGSLSAKQLTELAWGLEMSQQRFVWVVRPPVDGSACSAYLSANS GKIRDGTPDYLPEGFVSRTHERGFMVSSWAPQAEILAHQAVGGFLTHCGWNS ILESVVGGVPMIAWPLFAEQMMNATLLNEELGVAVRSKKLPSEGVITRAEIEA LVRKIMVEEEGAEMRKKIKKLKETAAESLSCDGGVAHESLSRIADESEHLLER VRCMARGA

Figure 3

ATGCATATCA CAAAACCACA CGCCGCCATG TTTTCCAGTC CCGGAATGGG CCATGICATC CCGGTGATCG AGCTTGGAAA GCGTCTCTCC GCTAACAACG GCTTCCACGT CACCGTCTTC GTCCTCGAAA CCGACGCAGC CTCCGCTCAA TCCAAGTTCC TAAACTCAAC CGGCGTCGAC ATCGTCAAAC TTCCATCGCC GGACATTTAT GGTTTAGTGG ACCCCGACGA CCATGTAGTG ACCAAGATCG GAGTCATTAT GCGTGCAGCA GTTCCAGCCC TCCGATCCAA GATCGCTGCC ATGCATCAAA AGCCAACGGC TCTGATCGTT GACTTGTTTG GCACAGATGC GTTATGTCTC GCAAAGGAAT TTAACATGTT GAGTTATGTG TTTATCCCTA CCAACGCACG TTTTCTCGGA GTTTCGATTT ATTATCCAAA TTTGGACAAA GATATCAAGG AAGAGCACAC AGTGCAAAGA AACCCACTCG CTATACCGGG GTGTGAACCG GTTAGGTTCG AAGATACTCT GGATGCATAT CTGGTTCCCG ACGAACCGGT GTACCGGGAT TTTGTTCGTC ATGGTCTGGC TTACCCAAAA GCCGATGGAA TTTTGGTAAA TACATGGGAA GAGATGGAGC CCAAATCATT GAAGTCCCTT CTAAACCCAA AGCTCTTGGG CCGGGTTGCT CGTGTACCGG TCTATCCAAT CGGTCCCTTA TGCAGACCGA TACAATCATC CGAAACCGAT CACCCGGTTT TGGATTGGTT AAACGAACAA CCGAACGAGT CGGTTCTCTA TATCTCCTTC GGGAGTGGTG GTTGTCTATC GGCGAAACAG TTAACTGAAT TGGCGTGGGG ACTCGAGCAG AGCCAGCAAC GGTTCGTATG GGTGGTTCGA CCACCGGTCG ACGGTTCGTG TTGTAGCGAG TATGTCTCGG CTAACGGTGG TGGAACCGAA GACAACACGC CAGAGTATCT ACCGGAAGGG TTCGTGAGTC GTACTAGTGA TAGAGGTTTC GTGGTCCCCT CATGGGCCCC ACAAGCTGAA ATCCTGTCCC ATCGGGCCGT TGGTGGGTTT TTGACCCATT GCGGTTGGAG CTCGACGTTG GAAAGCGTCG TTGGCGGCGT TCCGATGATC GCATGGCCAC TTTTTGCCGA GCAGAATATG AATGCGGCGT TGCTCAGCGA CGAACTGGGA ATCGCAGTCA GATTGGATGA TCCAAAGGAG GATATTTCTA GGTGGAAGAT TGAGGCGTTG GTGAGGAAGG TTATGACTGA GAAGGAAGGT GAAGCGATGA GAAGGAAAGT GAAGAAGTTG AGAGACTCGG CGGAGATGTC ACTGAGCATT GACGGTGGTG GTTTGGCGCA CGAGTCGCTT TGCAGAGTCA CCAAGGAGTG TCAACGGTTT TTGGAACGTG TCGTGGACTT GTCACGTGGT GCTTAG

MHITKPHAAM FSSPGMGHVI PVIELGKRLS ANNGFHVTVF VLETDAASAQ
SKFLNSTGVD IVKLPSPDIY GLVDPDDHVV TKIGVIMRAA VPALRSKIAA
MHQKPTALIV DLFGTDALCL AKEFNMLSYV FIPTNARFLG VSIYYPNLDK
DIKEEHTVQR NPLAIPGCEP VRFEDTLDAY LVPDEPVYRD FVRHGLAYPK
ADGILVNTWE EMEPKSLKSL LNPKLLGRVA RVPVYPIGPL CRPIQSSETD
HPVLDWLNEQ PNESVLYISF GSGGCLSAKQ LTELAWGLEQ SQQRFVWVVR
PPVDGSCCSE YVSANGGGTE DNTPEYLPEG FVSRTSDRGF VVPSWAPQAE
ILSHRAVGGF LTHCGWSSTL ESVVGGVPMI AWPLFAEQNM NAALLSDELG
IAVRLDDPKE DISRWKIEAL VRKVMTEKEG EAMRRKVKKL RDSAEMSLSI
DGGGLAHESL CRVTKECQRF LERVVDLSRG A

CCATGTCCTC CCGGTGATCG AGCTAGCTAA CCGCCGCGCGCGCCGCTCCGCCGCCGCCGCCCGCCCGCC	ATGCATATCA	CAAAACCACA	CGCCGCCATG	TTTTCCAGTC	CCGGAATGGG
TCCAAGCTCC TTAACTCAAC CGGTGTTGAC ATCGTCAACC TTCCATCGCGACATTTCT GGCTTGGTAG ACCCCAACGC CCATGTGGTA ACCAAGATTAACAACAACAACAACAACAACAACAACAACAACAACA	CCATGTCCTC	CCGGTGATCG	AGCTAGCTAA	GCGTCTCTCC	GCTAACCACG
CGACATTAT GGCTGAGAG CTCCAACGC CCATGTGGT ACCAGACT GAGTCATTAT GCGTGAAGCT GTTCCAACCC TCCGATCCAA GATCGTTG ATGCATCAAA ACCCAACGGC TCTGATCATT GACTTATGTC TTTATCGC CCAACGCGG TTATCTCGGA GTTTCGATCAT ATTATCCAAC TTTGGACG GTTATCAAAG AAGAGCACAC AGTGCAACGA AAACCGCTCA CTATACCG GTGTGAACCG GTTAGATTG AAGATATTAT GGATGCATAT CTGGTTCCC ACGAACCGGT GTACACGAT TTGGTTCGTC ACTGTCTGC CTACCCAA ACGAACCGGT GTACACGAT TTGGTTCGTC ACTGTCTGC CTACCCAA ACGAACCGGT GTACACGAA TACATGGGAA GAGATGGAGC CCAAATCAT GCGGATGGAA TCTTGGTGAA TACATGGGAA GAGATGGAGC CCAAATCAT AAAGTCCCTT CAAGACCCGA AACTTTTGGG CCGGGTCGCT CGTGTACCG CACCCGGTTT TTGATTGGTT AAACAAACAA CCAAACGAGT CGGTTCTC CATTTCCTTC GGGAGTGGTG GTTCTCTAAC GGCTCAACAG TTAACCGA TGGCGTGGGG GCTCGAGGAG AGCCAGCAAC GGTTTATATG GGTGGTTC CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAAGGC CTGAACCAAA GACAACACGC CAGAGTATCT ACCAGAACGG TTCGTGAGC ATCCTAGCCC ATCAGGCTCT TTGCAGTGAT TATTTCTCGG CTAAAGGC CTCGACGTTG GAAAGCATC CAGAGTATCT TCAAACAAT GCGTGGCCG TTTTCCCCGA GCAGAATATC ACCAGAAGGG TTCGTTGGC ATCCTAGCCC ATCAGGCCGT TGGTGGGTT TTAACCAAT GCGTGGCCG TTTTCCCCGA GCAGAATATG AACCAGAGG GCGATTTCTA GCGTGGCCG ATCCTAGCCC ATCAGGCGT TGCTGGCGT TCCAATGATA GCGTGGCCG TTTTCCCCCGA GCAGAATATG AACCGGCGCT TCCTAACGAT GAGCGCGC ATCCTGACGTTG GAAAGCGTCC TTTGCGGCGT TCCAATGATA GCGTGGCCG TTTTCCCCCGA GCAGAATATG AACCGCGCGT TGCTTAGCGA TGAACTGG ATCCTTGTTA GAGGAAGAT CACAAAGGAG GCGATTTCTA GGTCGAAG ATCCTCGTTA GAGGAAGGT TCCAAAGGAG GCGATTTCTA GGTCGAAG ATCCTCGTTA GAGGAAGGT TCCAAAGGAG GCGATTTCTA GGTCGAAG ATCCTCGTTA GAGGAAGGT TCCAAAGGAG GCGATTTCTA GGTCGAAG ATCCTCGTTA GAGGAAGGT TATGGCTGA GGACGAAGGT GAAGAGAT CAACGGTGGTG GTAGGAAGGT AAGAGAACACGG CGGAGATGT ACCTAACGGAG ATCCTGGTTG GAAGAAGTT AACACACGC CGGAGATGC ACTTAGTAC AACGGTGGTG GTAGGAAGGT AAGAGAACACGG CGGAGATGT ACCTTAGCGA ATCCTGGTGGTG TTTGGGGCGT TCCAAAGGAG GAAGAGAT ACCGGGGGTG GAAGAAGTT AACACACGC CGGAGATGC ACCTTAGTAC AACGGGGGGTG GAAGAAGTT AACACACGG CGGAGAGGT AACACGGG CGGAGAGGT GAAGAGATGCG CGGAGAGGT GAAGAGATGCACGG CGGAGAGGT GAAGAGATGCACGG CGGAGAGGT AACGGGGGGT TAGCGAGGT GAAGGAGAGGT GAAGAGATGCACGG CGGAGAGGT AACACGGGGGT TCCAAAGGAGG	GCTTCCACGT	CACCGTCTTC	GTCCTTGAAA	CTGACGCAGC	CTCCGTTCAG
ATCATCATA ACCCAACGC TCTGATCAT GACTTGTTG GCACAGATG GTTATGTCTT GCAGCGGAGT TAAACATGTT GACTTATGTC TTTATCGCC CCAACGCGCG TTATCTCGGA GTTTCGATAT ATTATCCAAC TTTAGCCG GTTATCAAAG AAGAGCACAC AGTGCAACGA AAACCGCTCA CTATACCGA GTGTGAACCG GTTAGATTG AAGATATTAT GGATGCATAT CTGGTTCG ACGAACCGGT GTACCACGAT TTGGTTCGTC ACTGTCTGGC CTACCCAA AAAGTCCCTT CAAGACCCGA AACTTTTGGG CCGGGTCGCT CGTGTACCGA AAAGTCCCTT CAAGACCCGA AACTTTTGGG CCGGGTCGCT CGTGTACCG CACCCGGTTT TTGATTGGTT AAACAAACAA CCAAACGAGT CGGTTCCC CACTCCGGTTT TTGATTGGT GTTCTTAAC GGCTCAACGG TTAAACCGA CCGCCCGTTT GCGCGGAGGAG AGCCAGCAAC GGTTTATATG GGTGGTTC CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAACGAC CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAACGAC CTCACACAGA GACAACACCC CAGAGTATCT ACCAGAACGG TTCGTGGC CTCACACAG TTGGTTCCAT CATGGGCAC GCAAGCTG CTCACCCGTTG TTGAGGGTTT TTAACCCAT GTGGTTGG CTCCACCGTTG GAAAGCGC TTGGTGGTTT TTAACCAAT GTGGTTGG CTCCACCGTTG GAAAGCGC TTGGTGGTTT TTAACCAAT GTGGTTGG CTCCACCGTTG GAAAGCGTC TTGCGGCGT TCCAATGAT GCGTGGGCC CTCGACGTTG GAAAGCGTC TTGCGGCGT TCCAATGAT GCGTGGCC CTCGACGTTG GAAAGCGTC TTGCGGCGT TCCAATGATA GCGTGGCC CTCGACGATG GAAAGCGTC TTGCGGCGT TCCAATGATA GCGTGGCC CTCGACGAGG GCAGAATATG AACCCGCGGT TGCTTAGCGA TGAACCGG CAAACGGAAAGTTC ACCAGAGGAG GAAGAGAT CAACGGCAAGGT GAAGAGATTC ACCAGAGGAG GAAGAGAT CAACGGCGAAGGT GAAGAGATTC ACCAGAGGAG GAAGAGAT CAACGGGAAAGGT GAAGAGATTC ACCAGAGGAG GAAGAGAT CAACGGGAAAGGT GAAGAGATTC ACCAGAGGAG GAAGAGAT CAACGGGAAAGGT GAAGAGATTC ACCAGAGGAGGT GAAGAGAT CAACGGGGGGGGT GAAGAGGT GAAGAGGT GAACCGG GGAGAGGT GAAGAGGAT CAACGGGGGGGGGGGGGGGGGGGGGGGGGG	TCCAAGCTCC	TTAACTCAAC	CGGTGTTGAC	ATCGTCAACC	TTCCATCGCC
ATGCATCAAA ACCCAACGGC TCTGATCATT GACTTGTTE GCACAGATT GTTATGTCTT GCAGCGGAGT TAAACATGTT GACTTATGTC TTTATCGC GTTATCAAAG AAGAGCACAC AGTGCAACGA AAACCGCTCA CTATACCGA GTGTGAACCG GTTAGATTG AAGATATTAT GGATGCATAT CTGGTTCC ACGAACCGGT GTACCACGAT TTGGTTCGTC ACTGTCTGGC CTAACCGAA AAAGTCCCTT CAAGACCCGA AACTTTTGGG CCGGGTCGCT CAAGACCGAT TTTATCCGGT TGGTCGTA TGCAGACCGA TACAATCATC CACGACCGA CACCCGGTTT TTGATTGGTT AAACAAACAA CCAAACGAGT CACGACCGA TGGCGTGGGG GCTCGAGGAG AGCCAGAACAA CCAAACGAGT CACGCCGAACGA TACAATCATC CACGACCGA CCGCCCGTTG ACGGCTGGT GTTCTCTAAC GGCTCAACAG TTAACCGAA CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAACGAG CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAACGAC GTACTTGCCAA ACCAACACGC CAGAGTATCT ACCAGAAGGG TTCGTGAC GTACTTGCGAA TACAACACGC CAGAGTATCT ACCAGAAGGG TTCGTGACC CTCCACCGTTG GAAACACCCC TTGCAGTGAT TATTTCTCGG CTAAACGAC CTCCACCGTTG GAAACACCCC TTGCAGTGAT TTAAACCAAT GTGGTTGG CTCCACCGTTG GAAACGCCCT TTGCAGCGT TCCAATGATA GCGTGGCCC CTCCACCGTTG GAAACGCCCT TTGCAGCGT TCCAATGATA GCGTGGCCC CTCCACCGTTG GAAACGCCC TTTGCGGCGT TCCAATGATA GCGTGGCCC CTCCACCGTTG GAAACGCTC TTTGCGGCGT TCCAATGATA GCGTGGCCC CTCGACCGTTG GAAACGTC TTTGCGGCGT TCCAATGATA GCGTGGCCC CTCGACCGTTG GAAACGTCC TTTGCGGCGT TCCAATGATA GCGTGGCCC CTCGACGTTG GAAACGTC CTTTGCGGCGT TCCAATGATA GCGTGGCCC CTCGACGTTG GAAACGTC CTTTGCGCGCT TCCAATGATA GCGTGGCCC CTCGACGTGG GAAACACGC CGGAAAGGT GAAACACGC CGGAAAGGT GAAAGAGT CAACGGTGGTG GAAAGATTC ACCACACGC CGGAAAGGT GAAACACGC CAAACCGC CAAACCTCC CACACCTACACACCCC CAAACCACGC CGGAAAGGT GAAACACGC CAAACCACCCCAAACGC CGGAAAGGT GAAACACGC CAAACCACCCCAAACGC CGGAAAGGT CACTTAGTACACACCGC CGAAACGCC CGAAACGCC CGAAACGCC CAAACGCACCCCAACGC CGAAACGCACCCACACCCACACCCACACCCACACCACCAC	CGACATTTCT	GGCTTGGTAG	ACCCCAACGC	CCATGTGGTG	ACCAAGATCG
CCAACGCGCG TTATCTCGGA GTTTCGATAT ATTATCCAC TTTGGACGCGCG TTATCTCGGA GTTTCGATAT ATTATCCAC TTTGGACGCGTTATCAAAG AAGAGCACAC AGTGCAACGA AAACCGCTCA CTATACCGCGCGTGTGAACCG GTTAGATTTG AAGATATTAT GGATGCATAT CTGGTTCCAACGAACCGAAC	GAGTCATTAT	GCGTGAAGCT	GTTCCAACCC	TCCGATCCAA	GATCGTTGCC
CCAACGCGC TTATCTCGGA GTTTCGATAT ATTATCCAAC TTTGGACGC GTTATCAAAG AAGACCACAC AGTGCAACGA AAACCGCTCA CTATACCGA GTGTGAACCG GTTAGATTTG AAGATATTAT GGATGCATAT CTGGTTCGC ACGAACCGGT GTACCACGAT TTGGTTCGTC ACTGTCTGGC CTACCCAAA AAAGTCCCTT CAAGACCCGA AACTTTTGGG CCGGGTCGCT CGTGTACCGAA AAAGTCCCTT CAAGACCCGA AACTTTTGGG CCGGGTCGCT CGTGTACCGAA CCACCGGTTT TTGATTGGTT AAACAAACAA CCAAACGAGT CGGTTCTC CACTTCCTTC GGGAGTGGTG GTTCTCTAAC GGCTCAACAG TTAACCGAA TGGCGTGGGG GCTCGAGGAG AGCCAGCAAC GGTTTATATG GGTGGTTC CCGCCCGTTG ACGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAAGGC GTACTTGCAAA GACAACACGC CAGAGTATCT ACCAGAACGAG TCGTGACC GTACTTGCGA TAGAGGTTTC ATGATCCCAT CATGGCCAC GCAAGCTG ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTTGG CTCGACGTTG GAAAGCGTC TTTGCGGCGT TCCAATGATA GCGTGGCC ATCTTCCTGCAA GCAGAATATG AACGCGCGT TCCAATGATA GCGTGGCC ATCTTCCTCTTA GAGGGATGA TCCAAAGGAG TGAACTGG ATCTCTGTTA GAGGGATGA TCCAAAGGAG GCGATTCTA GGGTCGAAGGA ATCCTTGTTA GAGGGATGA TCCAAAGGAG GCGATTCTA GGTCGAAGGA ATCCTTGTTA GAGGGAAGG TCCAAAGGAG GCGAATTCTA GGTCGAAGGA ATCTCTGTTA GAGGGAAGG TCCAAAGGAG GCGAATTCTA GGTCGAAGGA ATCCTCTGTTA GAGGGAAGG TCCAAAGGAG GCGAATTCTA GGTCGAAGGA ATCCTGGTTG GAAGAAGTTG AACGCGCGT TGCTTAGCGA TGAACTGG AAAGGAAAGG GCAAAAGGAG GCGAATGTC ACTTAGTAA ACCGGTGGTG GAAGAAGTTG AGAGACACGG CGGAAGGT GAAAGAGAT ACACGGTGGTG GTTCGGCGCA TGAACACGC CGGAAGGT ACTTAGTAA ACCGGTGGTG GAAGAAGTTG AGAGAACACGG CGGAAGGT ACTTAGTAA ACCGGTGGTG GAAGAAGTTG AGAGACACGG CGGAAGGT ACTTAGTAACACGGAAGGT GAAAGAGAT ACACGGTGGTG GTTCGGCGCA TGAACACGC CGGAAGGT ACTTAGTAACACGGAAGGT GAAAGAAGTTG AGAGACACGG CGGAAGGT ACTTAGTAACACGGAAGGT GAAAGAAGTTG AGAGACACGG CGGAAGGT ACTTAGACACGC CGGAAGGT GAAAGAGAT ACACGGTGGTG GTTCGGCCCA TGAGTCCCTT TGCAGAGGT GAAAGAGAT ACACGGTGGTG GTTCGGCCCA TGAGTCCCTT TGCAGAGGT GAAAGAGAT ACACGGAAAAGTTG AGAGACACGG CGGAAGGT CACTTAGTAACACCACGC CGGAAGGT TCCAAAGGAGGT GAAAGGAT ACACGGTGGTG GTTCGGCCGAAGGT TCCAAAGGAGGT GAAAGGAAGATCCACGC CGGAAGGT CACTTAGACACGC CGGAAGGT CACTTAGACACGC CGAAGGT CACTTAGACACGC CGGAAGGT CACTTAGACACGC CGGAAGGT CACTTAGACACCACGC CGGAAGGT CACTTAGACACGC CGAAGGT CACACGGT CACACGGT CACACGCTAA	ATGCATCAAA	ACCCAACGGC	TCTGATCATT	GACTTGTTTG	GCACAGATGC
GTTATCAAAG AAGAGCACAC AGTGCAACGA AAACCGCTCA CTATACCGC GTGTGAACCG GTTAGATTTG AAGATATTAT GGATGCATAT CTGGTTCC ACGAACCGGT GTACCACGAT TTGGTTCGTC ACTGTCTGGC CTACCCAA GCGGATGGAA TCTTGGTGAA TACATGGGAA GAGATGGAGC CCAAATCA AAAGTCCCTT CAAGACCCGA AACTTTTGGG CCGGGTCGCT CGTGTACCG CACCCGGTTT TTGATTGGTT AAACAAACAA CCAAACGAGT CGGTTCTC CATTTCCTTC GGGAGTGGTG GTTCTCTAAC GGCTCAACAG TTAACCGA TGGCGTGGGG GCTCGAGGAG AGCCAGCAAC GGTTTATATG GGTGGTTC CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAAGGA GTACTTGCGA TAGAGCTTC ATGATCCCAT CATGGGCACC GCAAGCTG ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTTGG CTCGACGTTG GAAGACGTCC TTTGCGGCGT TCCAATGATA GCGTGGCCC TTTTCCCCCGA GCAGAATATG AACGCGCGT TGCTTAGCGA TGAACTGG ATCCTTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAG ATCCTCGTTA GAGGGATGA TCCAAAGGAG GCGAATGTC GAAGAGAT TGAAGGCAACGC GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT CACGGTGGTG GTTCGGCCCA TGAGTCGCT TGCAAGGGT CACGAAGGAT CACGGTGGTG GTTCGGCCCA TGAGTCGCTT TGCAGAGGTC ACTTAGTAA CACGGTGGTG GTTCGGCCCA TGAGTCGCTT TGCAGAGGTC CACAAGGAG CACGGTGGTG GTTCGGCCCA TGAGTCGCTT TGCAGAGGTC CACAAGGAG CACGGTGGTG GTTCGGCCCA TGAGTCGCTT TGCAGAGGTC CACAAGGAG CACACGGTGGTG GTTCGGCCCA TGAGTCGCTT TGCAGAGGTC CACAAGGAG CACACGGTGGTG GTTCGGCCCA TGAGTCGCTT TGCAGAGGTC CACAAGGAG CACACGGTGGTG GTTCCGCCCA TGAGTCGCTT TGCAGAGGTC CACAAGGAG CACACGGTGGTG GTTCCGCCCA TGAGTCGCTT TGCAGAGGTC CACAAGGAG CACACGGTGGTG GTTCCGCCCA TGAGTCCCTT TGCAGAGGTCA CGAAGGAG CACACGGTGGTG GTTCCGCCCA TGAGTCCCTT TGCAGAGGTC CACAAGGAG CACACGGTGGTG GTTCGGCCCA TGAGTCCCTT TGCAGAGGTC CACAAGGAG CACACGGTGGTG GTTCGGCCCA TGAGTCCCTT TGCAGAGGTC CACAAGGAG CACACGGTGGTG GTTCGGCCCT TGAGCCCTT TGCAGAGGTC CACAAGGAG CACACGGTGGTC TTATGCCTGT TGCAGAGGTC CACAAGGAG CACACGGTGGTG GTTCGCTT TGCAGAGGTC CACAAGGAG CACACGGTGGTC TTATGCCTTA TGCACAGGTCA CGAAGAGTT CACACGAGTCA CGAGATGTC CACAGGAGTCA CGAAGGTCA CGAAGATACACGC CGAAGGTCA CGAAGATACACCGC CGAAGGTCA CGAAGATACACCGC CACACGTCACACCGC CGAAGAT	GTTATGTCTT	GCAGCGGAGT	TAAACATGTT	GACTTATGTC	TTTATCGCTT
GTGTGAACCG GTTAGATTTG AAGATATTAT GGATGCATAT CTGGTTCCCAAACGAACCGGT GTACCACGAT TTGGTTCGTC ACTGTTCGGC CTACCCAAACGAACCGGATGGAAACCGGATGGAAACCACGAAACCACGAACCCGAAACCACCGAAACCACC	CCAACGCGCG	TTATCTCGGA	GTTTCGATAT	ATTATCCAAC	TTTGGACGAA
ACGAACCGGT GTACCACGAT TTGGTTCGTC ACTGTCTGGC CTACCCAA GCGGATGGAA TCTTGGTGAA TACATGGGAA GAGATGGAGC CCAAATCA AAAGTCCCTT CAAGACCCGA AACTTTTGGG CCGGGTCGCT CACGACCG TTTATCCGGT TGGTCCGTTA TGCAGACCGA TACAATCATC CACGACCG CACCCGGTTT TTGATTGGTT AAACAAACAA CCAAACGAGT CGGTTCTC CATTTCCTTC GGGAGTGGTG GTTCTCTAAC GGCTCAACAG TTAACCGA TGGCGTGGGG GCTCGAGGAG AGCCAGCAAC GGTTTATATG GGTGGTTC CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAAGGC TGTAACCAAA GACAACACGC CAGAGTATCT ACCAGAAGGG TTCGTGAC GTACTTGCGA TAGAGGTTTC ATGATCCCAT CATGGGCACC GCAAGCTG CTCGACGTTG GAAAGCGTC TTGCGGCGT TCCAATGATA GCGTGGCC TTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTTAGCGA TGAACTGG ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGGTCGAACGG TGAGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT CAAGGGAAAGT GAAGAAGTTG AGAGACACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAGT CACGGTGGTG GTTCGGCCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAGT CACGGTGGTG GTTCGGCCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAGT CACGGTGGTG GTTCGGCCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAGA CACGGTGGTG GTTCGGCCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAGAGT CACGGTGGTG GTTCGCCGCTT TGCAGAGTCA CGAAGGAGT					
GCGGATGGAA TCTTGGTGAA TACATGGGAA GAGATGGAGC CCAAATCAC AAAGTCCCTT CAAGACCCGA AACTTTTGGG CCGGGTCGCT CGTGTACC TTTATCCGGT TGGTCCGTTA TGCAGACCGA TACAATCATC CACGACCG CACCCGGTTT TTGATTGGTT AAACAAACAA CCAAACGAGT CGGTTCTC CATTTCCTTC GGGAGTGGTG GTTCTCTAAC GGCTCAACAG TTAACCGAC TGGCGTGGGG GCTCGAGGAG AGCCAGCAAC GGTTTATATG GGTGGTTC CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAAGGC TGTAACCAAA GACAACACGC CAGAGTATCT ACCAGAAGGG TTCGTGAC GTACTTGCGA TAGAGGTTC ATGATCCCAT CATGGGCACC GCAAGCTG ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTTGG CTCGACGTTG GAAAGCGTCC TTTGCGGCGT TCCAATGATA GCGTGGCC TTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTTAGCGA TGAACTGG ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAG TGAGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT CAAGGGAAAGT GAAGAAGTTG AGAGACACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG	GTGTGAACCG	GTTAGATTTG	AAGATATTAT	GGATGCATAT	CTGGTTCCGG
AAAGTCCCTT CAAGACCCGA AACTTTTGGG CCGGGTCGCT CGTGTACCC TTTATCCGGT TGGTCCGTTA TGCAGACCGA TACAATCATC CACGACCGC CACCCGGTTT TTGATTGGTT AAACAAACAA CCAAACGAGT CGGTTCTC CATTTCCTTC GGGAGTGGTG GTTCTCTAAC GGCTCAACAG TTAACCGAC TGGCGTGGGG GCTCGAGGAG AGCCAGCAAC GGTTTATATG GGTGGTTC CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAAAGGC TGTAACCAAA GACAACACGC CAGAGTATCT ACCAGAAGGG TTCGTGAC GTACTTGCGA TAGAGGTTTC ATGATCCCAT CATGGGCACC GCAAGCTG ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTTGG CTCGACGTTG GAAAGCGTCC TTTGCGGCGT TCCAATGATA GCGTGGCC TTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTTAGCGA TGAACTGG ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAGGT TGAGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT CAAGGGAAAGT GAAGAAGTTG AGAGACACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG			•		-
TTTATCCGGT TGGTCCGTTA TGCAGACCGA TACAATCATC CACGACCGC CACCCGGTTT TTGATTGGTT AAACAAACAA CCAAACGAGT CGGTTCTC CATTTCCTTC GGGAGTGGTG GTTCTCTAAC GGCTCAACAG TTAACCGA TGGCGTGGGG GCTCGAGGAG AGCCAGCAAC GGTTTATATG GGTGGTTC CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAAGGC TGTAACCAAA GACAACACGC CAGAGTATCT ACCAGAAGGG TTCGTGAC GTACTTGCGA TAGAGGTTTC ATGATCCCAT CATGGCCACC GCAAGCTG ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTTGG CTCGACGTTG GAAAGCGTCC TTTGCCGCGT TCCAATGATA GCGTGGCC TTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTTAGCGA TGAACTGG ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAG TGAGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT CACGGTGGTG GTAGGAAGTTG AGAGACACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG					
CACCCGGTTT TTGATTGGTT AAACAACAA CCAAACGAGT CGGTTCTC CATTTCCTTC GGGAGTGGTG GTTCTCTAAC GGCTCAACAG TTAACCGA TGGCGTGGGG GCTCGAGGAG AGCCAGCAAC GGTTTATATG GGTGGTTC CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAAGGC TGTAACCAAA GACAACACGC CAGAGTATCT ACCAGAAGGG TTCGTGAC GTACTTGCGA TAGAGGTTTC ATGATCCCAT CATGGGCACC GCAAGCTG ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTTGG CTCGACGTTG GAAAGCGTCC TTTGCGGCGT TCCAATGATA GCGTGGCC TTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTTAGCGA TGAACTGG ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAG TGAGGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG				·	
CATTTCCTTC GGGAGTGGTG GTTCTCTAAC GGCTCAACAG TTAACCGA TGGCGTGGGG GCTCGAGGAG AGCCAGCAAC GGTTTATATG GGTGGTTC CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAAGGC TGTAACCAAA GACAACACGC CAGAGTATCT ACCAGAAGGG TTCGTGAC GTACTTGCGA TAGAGGTTTC ATGATCCCAT CATGGGCACC GCAAGCTG ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTTGG CTCGACGTTG GAAAGCGTCC TTTGCGGCGT TCCAATGATA GCGTGGCC TTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTTAGCGA TGAACTGG ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAG TGAGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CCTTAGTA					
TEGCCTTEG ACGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAAGGC TGTAACCAAA GACAACACGC CAGAGTATCT ACCAGAAGGG TTCGTGAC GTACTTGCGA TAGAGGTTTC ATGATCCCAT CATGGGCACC GCAAGCTG ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTTGG CTCGACGTTG GAAAGCGTCC TTTGCGGCGT TCCAATGATA GCGTGGCC TTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTTAGCGA TGAACTGG ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAG TGAGGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT GAAGGAAAGT GAAGAAGTTG AGAGACACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGGTC ACTTAGTA	CACCCGGTTT	TTGATTGGTT	AAACAAACAA	CCAAACGAGT	CGGTTCTCTA
CCGCCCGTTG ACGGCTCGTC TTGCAGTGAT TATTTCTCGG CTAAAGGCC TGTAACCAAA GACAACACGC CAGAGTATCT ACCAGAAGGG TTCGTGAC GTACTTGCGA TAGAGGTTTC ATGATCCCAT CATGGGCACC GCAAGCTG ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTTGG CTCGACGTTG GAAAGCGTCC TTTGCGGCGT TCCAATGATA GCGTGGCC TTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTTAGCGA TGAACTGG ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAG TGAGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT GAAGGAAAGT GAAGAAGTTG AGAGACACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG					
TGTAACCAAA GACAACACGC CAGAGTATCT ACCAGAAGGG TTCGTGACGGTACTTGCGA TAGAGGTTTC ATGATCCCAT CATGGGCACC GCAAGCTG ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTGGC TTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTAGCGA TGAACTGG ATCCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GAGTGGAAGG TCCAAAGGAG GCGATTTCTA GGTCGAAGGT GAAGAGATG GAAGAGATG AGAGGAAGGT GAAGAGATG CACGGTGGTG GAAGAAGTT AGAGCACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG CGAAGGT ACTTAGTA CACGGTGGTG GTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG CGAAGGAG CCGAAGGT ACTTAGTA CACGGTGGTG GTCCGCCCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG CCGAAGGT ACTTAGTA CACGGTGGTG GTCCGCCCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG CCGAAGGT ACCAAGGAG CCGAAGGT ACCAAGGAG CCGAAGGT ACCAAGGAG CCGAAGGT ACCAAGGAG CCGAAGGT ACCAAGGAG CCGAAGGT TGCAGAGTCA CGAAGGAG CCGAAGGT ACCAAGGAG CCAAGGT ACCAAGGAG CCAAGGT ACCAAGGAG CCAAGGT ACCAAGGAG CCGAAGGT ACCAAGGAG CCGAAGGT ACCAAGGAG CCGAAGGT ACCAAGGAG CCAAGGT ACCAAGGAG ACCACAG ACCAAGGAG ACCACAG ACCAAGGAG ACCACAG ACCAAGGAG ACCAAGGA ACCACACACA			•		
GTACTTGCGA TAGAGGTTTC ATGATCCCAT CATGGGCACC GCAAGCTG ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTTGG CTCGACGTTG GAAAGCGTCC TTTGCGGCGT TCCAATGATA GCGTGGCC TTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTTAGCGA TGAACTGG ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAG TGAGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT GAAGGAAAGT GAAGAAGTTG AGAGACACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG		•			
ATCCTAGCCC ATCAGGCCGT TGGTGGGTTT TTAACACATT GTGGTTGGCCCCCCCCCC					
CTCGACGTTG GAAAGCGTCC TTTGCGGCGT TCCAATGATA GCGTGGCCC TTTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTTAGCGA TGAACTGG ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAG TGAGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT GAAGGAAAGT GAAGAAGTTG AGAGCACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG	GTACTTGCGA	TAGAGGTTTC	ATGATCCCAT	CATGGGCACC	GCAAGCTGAA
TTTTCGCCGA GCAGAATATG AACGCGGCGT TGCTTAGCGA TGAACTGG ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAG TGAGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT GAAGGAAAGT GAAGAAGTTG AGAGCACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG					
ATCTCTGTTA GAGTGGATGA TCCAAAGGAG GCGATTTCTA GGTCGAAG TGAGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT GAAGGAAAGT GAAGAAGTTG AGAGACACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG	CTCGACGTTG	GAAAGCGTCC	TTTGCGGCGT	TCCAATGATA	GCGTGGCCGC
TGAGGCGATG GTGAGGAAGG TTATGGCTGA GGACGAAGGT GAAGAGAT GAAGGAAAGT GAAGAAGTTG AGAGACACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG					
GAAGGAAAGT GAAGAAGTTG AGAGACACGG CGGAGATGTC ACTTAGTA CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG	ATCTCTGTTA	GAGTGGATGA	TCCAAAGGAG	GCGATTTCTA	GGTCGAAGAT
CACGGTGGTG GTTCGGCGCA TGAGTCGCTT TGCAGAGTCA CGAAGGAG					
	GAAGGAAAGT	GAAGAAGTTG	AGAGACACGG	CGGAGATGTC	ACTTAGTATT
TCAACGGTTT TTGGAATGTG TCGGGGACTT GGGACGTGGT GCTTAG	CACGGTGGTG	GTTCGGCGCA	TGAGTCGCTT	TGCAGAGTCA	CGAAGGAGTG
100000000	TCAACGGTTT	TTGGAATGTG	TCGGGGACTT	GGGACGTGGT	GCTTAG

MHITKPHAAM FSSPGMGHVL PVIELAKRLS ANHGFHVTVF VLETDAASVQ
SKLLNSTGVD IVNLPSPDIS GLVDPNAHVV TKIGVIMREA VPTLRSKIVA
MHQNPTALII DLFGTDALCL AAELNMLTYV FIASNARYLG VSIYYPTLDE
VIKEEHTVQR KPLTIPGCEP VRFEDIMDAY LVPDEPVYHD LVRHCLAYPK
ADGILVNTWE EMEPKSLKSL QDPKLLGRVA RVPVYPVGPL CRPIQSSTTD
HPVFDWLNKQ PNESVLYISF GSGGSLTAQQ LTELAWGLEE SQQRFIWVVR
PPVDGSSCSD YFSAKGGVTK DNTPEYLPEG FVTRTCDRGF MIPSWAPQAE
ILAHQAVGGF LTHCGWSSTL ESVLCGVPMI AWPLFAEQNM NAALLSDELG
ISVRVDDPKE AISRSKIEAM VRKVMAEDEG EEMRRKVKKL RDTAEMSLSI
HGGGSAHESL CRVTKECQRF LECVGDLGRG A

p-coumaryl alcohol

Sinapyl alcohol

p-coumaryl aldehyde

Coniferyl aldehyde

Sinapyl aldehyde

*: position for glucosylation

Figure 8a

ATGAAGATTACAAAACCACATGTGGCCATGTTCGCTAGCCCCGGAATGGGCCACATC
ATCCCGGTGATCGAGCTCGGAAAACGCTTAGCTGGTTCCCACGGCTTCGATGTCACC
ATTTTCGTCCTTGAAACCGACGCAGCCTCAGCTCAATCTCAATTCCTTAACTCACCA
GGCTGCGACGCGGCCCTTGTTGATATCGTTGGCCTCCCAACGCCCGATATCTCCGGT
TTAGTCGACCCATCAGCCTT

Figure 8b

TGTGGTGACCAAGATCGGAGTCATTATGCGTGAAGCTGTTCCAACCCTCCGATCCAA GATCGTTGCCATGCATCAAAACCCAACGGCTCTGATCATTGACTTGTTTTGGCACAGA TGCGTTATGTCTTGCAGCGGAGTTAAACATGTTGACTTATGTCTTTATCGCTTCCAA CGCGCGTTATCTCGGAGTTTCGATATATTATCCAACTTTGGACGAAGTTATCAAAGA AGAGCA

Figure 8c

CACAGTGCAAAGAACCCACTCGCTATACCGGGGTGTGAACCGGTTAGGTTCGAAGA TACTCTGGATGCATATCTGGTTCCCGACGAACCGGTGTACCGGGATTTTGTTCGTCA TGGTCTGGCTTACCCAAAAGCCGATGGAATTTTGGTAAATACATGGGAAGAGATGGA GCCCAAATCATTGAAGTCCCTTCTAAACCCAAAGCTCTTGGGCCGGGTTGCTCGTGT ACCGGTCTATCCAATCGGT